Measures for Determining English Language Proficiency and the Resulting Implications for Instructional Provision and Intervention

Craig A. Albers, Dorry M. Kenyon and Timothy J. Boals

Assessment for Effective Intervention 2009; 34; 74 originally published online May 5, 2008;
DOI: 10.1177/1534508408314175

Assessment for Effective Intervention
Volume 34 Number 2
March 2009 74-85

10.1177/1534508408314175
http://aei.sagepub.com
hosted at
http://online.sagepub.com

Measures for Determining English Language Proficiency and the Resulting Implications for Instructional Provision and Intervention

Craig A. Albers
University of Wisconsin–Madison
Dorry M. Kenyon
Center for Applied Linguistics, Washington, D.C.
Timothy J. Boals
Wisconsin Center for Education Research, Madison

Although numerous English language proficiency (ELP) measures currently exist, many were developed prior to the No Child Left Behind Act of 2001 (NCLB). These pre-NCLB measures typically focused on social language proficiency, whereas post-NCLB measures are linked to ELP standards and focus on academic language proficiency (ALP). ELP measures are typically used for accountability purposes and to determine eligibility for services; less attention has been given to their utility in enhancing classroom instruction and intervention provision. Inconsistency in scores between pre- and post-NCLB measures frequently leaves educators wondering whether English language learners (ELLs) have the necessary ALP to benefit from classroom instruction. This study investigates the intervention validity of ELP assessment by examining the concurrent validity of various pre-NCLB measures to a recently developed post-NCLB measure. As hypothesized, results indicate moderate correlations between pre- and post-NCLB measures, suggesting that ALP-focused post-NCLB measures are likely to provide more utility for ELL classroom instruction and intervention provision.

Keywords: *English language learners; English language proficiency; utility; intervention*

Increasing numbers of students classified as English language learners (ELLs), combined with increased accountability requirements as mandated by the No Child Left Behind Act of 2001 (NCLB), have led to an increased interest regarding the measurement of English language proficiency (ELP). Current ELP measures, which were designed to meet NCLB accountability requirements and are to be linked directly to state ELP standards, also have been proposed to be beneficial in guiding classroom instruction and intervention. However, the evidence base for the instructional utility of ELP measures has yet to be established. Therefore, we attempt to contribute to this utility evidence base by examining issues of concurrent validity among a wide variety of ELP measures that are currently available. Establishing this concurrent validity is significant, as the high mobility rates of language-diverse students result in situations where ELL students enroll in a different school and arrive with a variety of ELP scores, some of which may not be acceptable to the school in which the student has enrolled. If scores from varying ELP measures indicate differing ELP levels, the school is then unable to determine which ELP score is more indicative of the student's skills and understanding, thereby limiting possible intervention validity at the classroom (i.e., instruction) and individual levels. To address the intervention validity of ELP measures, we (a) examine the current status of ELL students in U.S. schools, (b) provide a review of ELP accountability requirements and ELP measures, (c) discuss possible uses of ELP assessment results, and (d) provide data from commonly used ELP measures to examine concurrent validity.

Status of English Language Learners in Public Schools

The number of ELLs has been increasing at a significant rate; for example, approximately 50% of the 10 million children ages 5 to 17 years who speak a language

Editor's Note: This article should be considered as part of AEI 34(1), December 2008, Special Issue: Intervention Validity of Assessments Used in Special Education (Andrew T. Roach, Guest Editor).

other than English at home are classified as ELLs in the school setting, accounting for 10% of the total student population—an increase of over 80% in the past decade (Gottlieb, 2006; National Center for Education Statistics, 2007). More than one half of the nation's public schools have at least one student classified as an ELL enrolled in the school. Furthermore, it is estimated that by the year 2030, 40% of the school population will speak English as a second language (U.S. Department of Education and National Institute of Child Health and Human Development, 2003).

By nature of their second language status, students classified as ELLs generally are considered to be at risk for academic failure and poor developmental outcomes, both in the short and long terms. ELL students consistently lag behind their peers in academic achievement and experience higher rates of negative occupational, economic, social, and physical outcomes. Although these disparities are likely because of a variety of reasons, it is clear that these students are frequently denied similar opportunities to learn the necessary academic skills and tend to be placed in remedial programs at the expense of experiences with typical grade-level content.

English Language Proficiency

The seminal work of Cummins (1980, 1981) illustrated that the acquisition of language proficiency is a not dichotomous process but rather one that develops along a continuum. Social English proficiency (i.e., basic interpersonal conversational skills) takes approximately 2 to 3 years to develop, whereas Academic English proficiency (i.e., cognitive academic language proficiency) takes approximately 5 to 7 years to develop. Academic English proficiency has been found to be essential and it may take significantly longer for a student to develop the academic language proficiency needed for school success (Cummins, 1984). In addition to learning academic subjects in school, ELLs are confronted with an additional learning task—learning the English language—referred to as the curriculum's *double demand* (Baker, Kameenui, & Simmons, 2002; Baker, Plasecia-Peinado, & Lezcano-Lytle, 1998; Gersten, 1999). Moreover, the adequate conversational English of some ELL students falsely leads some educators to believe that these students' language skills are sufficient for completion of academic tasks (Gravois & Gickling, 2002).

Legal Requirements Pertaining to ELP Assessment

Although federal, state, and district policies vary in their definitions of ELLs, ELLs are generally regarded as "linguistically and culturally diverse students whose lack of English language proficiency precludes them from accessing, processing, or learning grade-level material in English" (Gottlieb, 2006, p. 186). Because school districts are required under federal civil rights law to assist students who are ELLs in overcoming language barriers so they can meaningfully participate in educational programs (Office for Civil Rights, U.S. Department of Education, 2000), districts need to systematically identify all students who are potential ELLs.

Title I of NCLB requires that states "provide for an annual assessment of English proficiency (measuring students' oral language, reading, and writing skills in English) of all students with limited English proficiency in schools served by the State educational agency" (sec. 1111(b)(7)). Title I requires states using federal funds to assess non-English-speaking students' English proficiency for the purpose of demonstrating adequate yearly progress (AYP; U.S. Department of Education, 2004). The ELP assessments must measure students' reading, comprehension, writing, speaking, and listening language proficiency (U.S. Department of Education, 2004). These assessments must also be aligned with state ELP standards (Lara, 2005). Title III of NCLB is similar to Title I in that it requires states to establish English proficiency standards, implement annual ELP assessments, and define annual objectives for increasing and measuring children's development and attainment of English proficiency.

Prior to the passage of NCLB (2001), states were not required by federal law to have ELP standards or uniform, reliable assessments for ELLs to measure their progress toward the attainment of academic ELP. The typical standards consisted of a locally devised curriculum, which could range from exceptional to seriously inadequate. The shift in paradigms from pre-NCLB approaches centered on grammar, pronunciation, and language basics to post-NCLB approaches that focus on content-based methodologies that teach language through, and simultaneously with, grade level academic content has served to encourage more appropriate instruction and goals.

Current ELP Measures Used by States

Prior to 2001, there were a variety of commercially marketed ELP assessments available. Because these ELP assessments were designed before the passage of NCLB, many of these instruments are questionable in meeting the requirements set forth in NCLB. There is substantial disparity among tests of ELP in terms of their theoretical orientation, design, materials, and reporting of results. Furthermore, there are significant questions regarding how these measures provide guidance to educators for classroom instruction and corresponding intervention

Table 1
English Language Proficiency Measures Currently Being Used by States for Accountability Requirements

ELP Measure	States
ACCESS for ELLs	Alabama, Delaware, District of Columbia, Georgia, Illinois, Kentucky, Maine, New Hampshire, New Jersey, North Dakota, Oklahoma, Pennsylvania, Rhode Island, Vermont, Virginia (beginning in 2009), Wisconsin
CELLA	Florida
ELDA	Arkansas, Iowa, Louisiana, Nebraska, South Carolina, Tennessee, West Virginia
ELPA	Michigan, Nevada, Oregon
IPT	Alaska, North Carolina
LAS Links	Connecticut, Hawaii, Indiana, Maryland,
MAC-II	Missouri
SELP	Mississippi, Virginia
State-specific ELP measure	Arizona, California, Colorado, Idaho, Kansas, Massachusetts, Minnesota, Montana, New Mexico, New York, Ohio, South Dakota, Texas, Utah, Washington, Wyoming

Note: ELP = English language proficiency; ACCESS for ELLs = *Assessing Comprehension and Communication in English State-to-State for English Language Learners*; CELLA = *Comprehensive English Language Learning Assessment*; ELDA = *English Language Development Assessment*; ELPA = *English Language Proficiency Assessment*; IPT = *IDEA Proficiency Test*; LAS Links = *Language Assessment System Links*; MAC-II = *Maculaitis Assessment of Competencies Test of English Language Proficiency*; SELP = *Stanford English Proficiency Test*.

Table 2
Comparison of Features of the Pre-NCLB and Post-NCLB ELP Measures

Pre-NCLB ELP Measures	Post-NCLB ELP Measures
Not compliant with the requirements of NCLB	Compliant with the requirements of NCLB
Not based on ELP standards	Based on ELP standards (e.g., consortium, state, organization, etc.)
Emphasis on social language	Emphasis on academic language
Generally integrated oral language domains	Independent language domains (i.e., listening and speaking)
Limited utility for classroom instruction and intervention	Increased utility for classroom instruction and intervention
Different tests for each grade-level cluster (no comparability)	Vertically scaled across tiers and grade-level clusters
Low-stakes outcomes	High-stakes outcomes

Note: NCLB = No Child Left Behind Act of 2001; ELP = English language proficiency.

provision. Commonly used ELP measures, both pre- and post-NCLB passage, are described briefly below. In addition, Table 1 indicates which ELP measures are currently being used for NCLB accountability purposes within the United States. Finally, Table 2 compares features of the pre-NCLB ELP measures to the Assessing Comprehension and Communication in English State-to-State for English Language Learners (ACCESS for ELLs), which is classified as a post-NCLB ELP measure.

Idea Proficiency Test (IPT). The IPT is an ELP assessment designed to assess and categorize students as nonlimited, or fluent English speaking, reading, or writing. The IPT was originally developed in 1979 and revised in 1991. Results are reported as three designations (six levels of oral language proficiency) and three levels of literacy. A post-NCLB revised version of the IPT is currently used in Alaska and North Carolina.

Language Assessment Scales (LAS). The LAS was originally developed in response to the Supreme Court decision in *Lau v. Nichols* (1974). The most current literacy sections date back to 1988, the oral language sections were last revised in 1990, and the Pre-LAS was introduced in 1998. Speaking and writing sections utilize rubrics for teacher scoring, whereas listening and reading are selected response. Tapes are available for administration of the listening and speaking sections. The LAS yields five levels of oral language proficiency and three levels of literacy. A post-NCLB revised version of the LAS, the LAS Links, is currently used in Connecticut, Hawaii, Indiana, and Maryland.

Language Proficiency Test Series (LPTS). Launched in 1999, the LPTS's conceptual base and design were derived from the Illinois Measure of Annual Growth in English, which served as the state's ELP test. Unlike the other language proficiency measures, LPTS draws from

themes with sets of interrelated items. The LPTS provides a continuous progression of scale scores from age 3 years through Grade 12, so scores can be compared from one grade level to the next. The results are reported as two levels of oral language proficiency and four levels of literacy. No states currently use the LPTS for accountability purposes.

Maculaitis Assessment of Competencies Test of English Language Proficiency (MAC-II). In its second edition, the MAC-II has the most grade-level cluster forms (five) of the four ELP tests that measure oral language and literacy. Listening is assessed and reported independently from speaking. Five-point rubrics are used to interpret oral language and writing samples. Results are converted into five levels of oral language competency and five levels of literacy. A post-NCLB revised MAC-II is currently being used by Missouri.

Stanford English Language Proficiency Measure (SELP). The SELP is the most recent commercially designed ELP instrument and was designed following the passage of NCLB. The SELP provides results in Listening, Writing Conventions, Reading, and Speaking scores. Student performance is classified as being at the emergent, basic, intermediate, or proficient levels. The SELP is currently being used by Mississippi and Virginia.

State-specific ELP measures. For a variety of reasons, 17 states have independently developed their own ELP measures that are directly linked to their ELP state standards. Readers interested in these measures are encouraged to contact their respective state educational agency (SEA) for more information regarding their ELP measures.

Consortium-based ELP measures. Following the passage of NCLB, four independent consortia of states were formed to address ELL-related issues, including the assessment of ELP (Abedi, 2007). These four—the Comprehensive English Language Learning Assessment Consortium, the English Language Development Assessment Consortium, Mountain West Consortium, and the World-Class Instructional Design and Assessment (WIDA) Consortium—have each developed separate ELP assessments that are directly linked to ELP standards.

The WIDA Consortium (www.wida.us), currently the largest ELL consortium, consists of 16 states and accounts for more than 420,000 ELLs in kindergarten through Grade 12 in more than 3,500 school districts. WIDA (2004; 2007) has created and adopted comprehensive ELP Standards that address the need for students to become fully proficient in *both* social and academic English; student ELP levels are obtained through administration of the WIDA ELP measure, titled the ACCESS for ELLs. The consortium first developed the WIDA ELP Standards to provide a framework for PreK through Grade 12 ELL instruction and assessment. The WIDA ELP Standards include both a model for assessing and promoting English language development in the classroom and a model for large-scale assessment. There are five fundamental ELP Standards (WIDA, 2004, p. 3):

1. ELLs communicate for *social and instructional* purposes within the school setting;
2. ELLs communicate information, ideas, and concepts necessary for academic success in the content area of *language arts*;
3. ELLs communicate information, ideas, and concepts necessary for academic success in the content area of *mathematics*;
4. ELLs communicate information, ideas, and concepts necessary for academic success in the content area of *science*; and
5. ELLs communicate information, ideas, and concepts necessary for academic success in the content area of *social studies*.

These new ELP Standards and their accompanying model performance indicators (MPIs) provide illustrations of student expectations at each proficiency level and contain information about the language function, content, and type of support associated with a particular task. More than 800 total MPIs, designed to assist educators in classroom and curriculum planning and also functioning as an assessment framework, are provided across the language proficiency levels (WIDA, 2004).

Uses and Utility of ELP Measures

ELP measures and resulting scores can be used for a variety of purposes, including (a) determination of eligibility for services, (b) meeting NCLB accountability requirements, (c) guiding classroom instruction, and (d) assisting in intervention design and implementation.

Eligibility

ELP measures can be used as one component of a comprehensive assessment system to determine eligibility for ELL-related services. For example, when a new student enrolls in a school, a home language survey is completed by the student's parent or parents. If the home language survey indicates that the student uses a language other than English at home, or if a language other

than English is used by an individual or individuals in the home, the student typically completes an ELP measure. If the results of this measure indicate that the student meets the requirements as an ELL, the student is then eligible to receive ELL-related services offered by the school and the LEA.

ELP measures can also be utilized as one component of an individualized comprehensive assessment to determine eligibility for special education services. As has been extensively documented elsewhere (e.g., Abedi, 2006; Laing & Kamhi, 2003; McCardle, Mele-McCarthy, Cutting, Leos, & D-Emilio, 2005), disentangling issues between ELP and academic difficulties is an extremely complex issue that has yet to be adequately resolved. However, the connection between lower levels of ELP and academic achievement clearly indicates that ELP is highly correlated with academic performance. Consequently, an ELL student at a lower ELP level is likely to have difficulties with academic content; however, this should not be automatically seen as an indicator of a specific learning disability or other underlying disability. Thus, when concerns about an ELL student are brought to the attention of educators, consideration of the student's ELP level should be considered. The potential drawback of such an approach is that early intervention services might be delayed while trying to determine the relative contribution of lower ELP levels on the student's academic performance; data suggest that schools are identifying learning difficulties in the ELL population at increased rates in Grades 4 through 6, basically 2 to 3 years after they are typically detecting similar difficulties in monolingual English students (McCardle et al., 2005). These data highlight the difficulty educators have sorting out issues of language acquisition from skill deficits commonly associated with learning difficulties.

Accountability

NCLB Title III requires SEAs to (a) establish ELP standards, (b) conduct an annual standards-based assessment of ELP, and (c) determine AYP and annual measurable achievement objectives for meeting state academic standards and increasing the percentage of ELLs progressing and attaining ELP, respectively. Under Title I, SEAs must determine AYP in the academic content areas of reading or language arts, math, and science. Title III requires that SEAs, LEAs, and schools include data regarding the percentage of ELL students making progress toward ELP and the percentage of ELL students who have attained ELP. Because of these requirements, it is essential that ELP measures be psychometrically sound *and* have intervention validity. For an ELP measure to have intervention validity, it must provide information that can assist educators in (a) guiding classroom instruction, (b) providing additional interventions, and (c) monitoring outcomes (Gresham, 2007). Failure to consider the consequential and intervention validity in the design and administration of ELL measures raises the very real possibility that inadequate measures are being used to guide important educational decisions, to the detriment of ELL students.

Instruction

Although most states had requirements to identify and assess ELLs prior to NCLB, the type of assessment used was often open to the discretion of local districts. The typical ELP test was administered and scored at the local level. These assessments did an adequate to good job of identifying ELLs for program support. They had not been designed, however, with the standards-based movement or NCLB higher stakes requirements in mind. Hence, they did not provide consistent, reliable scores across time and grade levels. The tests were to a large extent modeled from an earlier conceptualization of English skills that emphasized social or conversational contexts rather than academically based language specifically needed to perform in classrooms. This meant that the diagnostic information available for teachers was usually inadequate for fine-tuning support, determining when ELLs were ready to "exit" from program services, or judging the efficiency of one programmatic approach as opposed to another. Because states often did not require uniformity in the selection of ELP assessments prior to NCLB, comparisons among schools, districts, or states were hampered (August & Hakuta, 1998; Echevarria, Vogt, & Short, 2004; Gottlieb, 2006; Scarcella, 2003)

Intervention

As discussed in other articles of this special issue, the concept of the intervention validity has been raised as a framework to understand how assessment is linked to interventions (Hayes, Nelson, & Jarrett, 1987; Yoon & Resnick, 1998). Assessment is considered to have intervention validity if some positive outcome occurs during the instruction or intervention process, or on the termination of intervention, as influenced by the assessment device, distinction, or strategy. Within this definition, a wide variety of assessments may contribute to treatment selection or monitoring but may have no relationship to positive treatment outcomes. Therefore, *utility* emphasizes treatment outcomes, not merely treatment decisions. Within a treatment validity framework (e.g., Fuchs,

Table 3
Participants by Grade and ELP Measure

	ELP Measure				
Grade	IPT	LAS	LPTS	MAC-II	Total[a]
K	102	81	246	47	476
1	109	184	216	95	604
2	143	137	246	76	602
3	102	80	290	63	535
4	82	57	146	74	359
5	104	32	216	57	409
6	116	55	142	97	410
7	111	110	58	110	389
8	106	62	48	142	358
9	28	12	150	134	324
10	37	17	120	106	280
11	30	2	92	79	203
12	9	2	31	43	85
Total *N*	1,079	831	1,952	1,123	4,985

Note: ELP = English language proficiency; IPT = *IDEA Proficiency Test*; LAS = *Language Assessment System*; LPTS = *Language Proficiency Test Series*; MAC-II = *Maculaitis Assessment of Competencies Test of English Language Proficiency*.
a. All students completed the *Assessing Comprehension and Communication in English State-to-State for English Language Learners*.

Fuchs, & Speece, 2002), an ELP assessment would have intervention validity if the student were making or beginning to make adequate progress as a result of information obtained by the ELP measure.

Rationale for Study

Drawing on data from a bridge study conducted by the WIDA Consortium in 2005 (Gottlieb & Kenyon, 2006), the current study was designed to examine the strength of the relationship between the results from four older versions of ELP tests that were commonly used pre-NCLB and the results from a post-NCLB ELP measure, the ACCESS for ELLs. Thus, the goal was to establish links between older measures of ELP and a new measure of ELP when a representative sample of ELLs and proficient English students are double tested within a specified time frame. It was hypothesized that moderate correlations would be obtained between the ACCESS for ELLs and the pre-NCLB measures, as each was designed to be an ELP measure. However, the ACCESS for ELLs was designed to meet NCLB requirements and assess the ELP needed for students to academically succeed in the classroom; thus, correlations should not be very strong, as strong correlations would indicate that the post-NCLB measure is not necessary. Conversely, low correlations would be interpreted as indicating that the ELP measures are examining different constructs.

Method

Participants

Participants consisted of 4,985 students in Grades K through 12 from selected districts in WIDA Consortium states. Table 3 shows the number of students who participated by test and by grade level.

Measures and Procedure

All students completed the ACCESS for ELLs and one additional ELP measure. For the second ELP measure, students completed the IPT, LAS, LPTS, or MAC-II (measures described in more detail above), which were the measures that were previously administered in the students' districts prior to the passage of NCLB legislation. Within 3 months of taking the ACCESS for ELLs, these students were then administered one of four other tests used by the district, consisting of the IPT, LAS, LPTS, or MAC-II.

Results

Correlation coefficients were computed between the ACCESS for ELLs and the (a) IPT, (b) LAS, (c) LPTS, and (d) MAC-II subtests. Results are disaggregated below for each set of analyses. Consistent with Cohen's (1988) recommendations, correlations of .10, .30, and .50 were interpreted as small, medium, and large, respectively.

Table 4
Correlations Between Subtest Scores on the IPT and Subtest Scores on the ACCESS for ELLs by Grade-Level Cluster and Combined Across Clusters

	ACCESS for ELLs Grade-Level Cluster											
IPT Form	K		1–2		3–5		6–8		9–12		Combined	
ACCESS reading												
Reading												
EL	.572**	(102)	.543**	(103)	—		—		—		.741**	(205)
1A, 1B	—		.440**	(149)	.682**	(101)	—		—		.540**	(250)
2A, 2B	—		—		.659**	(184)	.539**	(112)	—		.618**	(296)
3A, 3B	—		—		—		.739**	(213)	.694**	(104)	.713**	(317)
ACCESS writing												
Writing												
EL	.165	(102)	.531**	(103)	—		—		—			
1A, 1B	—		.519**	(149)	.716**	(101)	—		—		.566**	(250)
2A, 2B	—		—		.570**	(186)	.532**	(113)	—		.550**	(299)
3A, 3B	—		—		—		.731**	(211)	.482**	(100)	.631**	(311)
ACCESS listening												
Oral[a]												
1C, 1D	.336	(23)	.541**	(79)	.748**	(91)	.606**	(26)	—		.710**	(219)
1E, 1F	.503**	(71)	.547**	(147)	.323**	(141)	.639	(10)	—		.578**	(369)
2C, 2D	—		—		—		.627**	(84)	.523**	(83)	.515**	(167)
ACCESS speaking												
Oral[a]												
1C, 1D	.615*	(23)	.425**	(79)	.732**	(92)	.568*	(26)	—		.594**	(220)
1E, 1F	.556**	(71)	.494**	(147)	.605**	(141)	.746*	(10)	—		.603**	(369)
2C, 2D	—		—		—		.729**	(84)	.696**	(81)	.678**	(165)

Note: IPT = *IDEA Proficiency Test*; ACCESS for ELLs = *Assessing Comprehension and Communication in English State-to-State for English Language Learners*. Values in parentheses are *n* values for the correlation analysis.
a. The IPT does not have separate Listening and Speaking scores.
*$p < .01$, two-tailed. **$p < .001$, two-tailed.

Correlations Between IPT and ACCESS for ELLs

The IPT does not provide separate scores for Listening and Speaking, instead providing a single Oral Language score. Consequently, the IPT Oral Language scores were compared with the ACCESS for ELLs Listening and Speaking subtests. Table 4 provides all of the obtained correlations. As indicated, all correlations between the subtests of the IPT and the ACCESS for ELLs were statistically significant ($p < .01$ or $p < .001$) and greater than .42, except for the correlations between Writing subtests for the kindergarten grade-level cluster (.17) and the IPT Oral Language and the ACCESS for ELLs Listening subtest (.34). Numerous correlations exceeded .70, including the IPT Reading and ACCESS for ELLs Reading across the Grades 6 through 8 cluster, between the IPT and ACCESS for ELLs Writing subtests in the Grades 3 through 5 and Grades 6 through 8 clusters, between the IPT Oral and ACCESS for ELLs Listening subtests across the Grades 3 through 5 cluster, and between the IPT Oral and ACCESS for ELLs Speaking subtests across the Grades 3 through 5 and Grades 6 through 8 clusters.

Correlations Between LAS and ACCESS for ELLs

Similar to the IPT, the LAS provides a single Oral Language score as compared to separate scores for Listening and Speaking, as with the ACCESS for ELLs. The obtained correlations between the LAS and ACCESS for ELLs are listed in Table 5. All of the correlations between subtests were statistically significant ($p < .01$ or $p < .001$) and greater than .32. Correlations exceeding .70 included Reading across the Grades 1 through 2 and 3 through 5 clusters and LAS Oral Language and ACCESS for ELLs Listening subtests at the Grades 1 through 2 cluster.

Correlations Between LPTS and ACCESS for ELLs

The LPTS provides a single Language and Speaking score, which resulted in statistically significant correlations

Table 5
Correlations Between Subtest Scores on the LAS and Subtest Scores on the ACCESS for ELLs by Grade-Level Cluster and Combined Across Clusters

	ACCESS for ELLs Grade-Level Cluster											
LAS Form	K		1–2		3–5		6–8		9–12		Combined	
ACCESS reading												
Reading												
Pre-LAS preliteracy	.376**	(79)	.557**	(50)	—		—		—		.317**	(129)
R/W 1A, 1B	—		.754**	(159)	.731**	(80)	—		—		.757**	(239)
R/W 2A, 2B	—		—		.683**	(84)	.616**	(53)	—		.668**	(137)
R/W 3A, 3B	—		—		—		.600**	(168)	.710**	(33)	.633**	(201)
ACCESS writing												
Writing												
Pre-LAS preliteracy	.493**	(79)	.387*	(50)	—		—		—		.323**	(129)
R/W 1A, 1B	—		.561**	(161)	.798**	(80)	—		—		.684**	(240)
R/W 2A, 2B	—		—		.567**	(85)	.744**	(53)	—		.587**	(139)
R/W 3A, 3B	—		—		—		.584**	(168)	.762**	(33)	.507**	(201)
ACCESS listening												
Oral[a]												
Pre-LAS oral	.540**	(81)	.733**	(20)	—		—		—		.474**	(101)
1C, 1D	—		.451**	(250)	.562**	(134)	.484**	(48)	—		.525**	(432)
2C, 2D	—		—		—		.482**	(165)	—		.509**	(198)
ACCESS speaking												
Oral[a]												
Pre-LAS oral	.547**	(81)	.539*	(21)	—		—		—		.548**	(102)
1C, 1D	—		.505**	(250)	.595**	(134)	.533**	(48)	—		.562**	(432)
2C, 2D	—		—		—		.595**	(165)	.672**	(33)	.599**	(198)

Note: LAS = *Language Assessment System*; ACCESS for ELLs = *Assessing Comprehension and Communication in English State-to-State for English Language Learners*; R/W = Reading/Writing. Values in parentheses are *n* values for the correlation analysis.
a. The LAS does not have separate Listening and Speaking scores.
*$p < .01$, two-tailed. **$p < .001$, two-tailed.

($p < .01$ or $p < .001$) with the ACCESS for ELLs Listening and Speaking subtests. All correlations between the LPTS and ACCESS for ELLs Reading and Writing subtests were also statistically significant and greater than .48. Specific correlations are listed in Table 6.

Correlations Between MAC-II and ACCESS for ELLs

Correlations between subtests on the MAC-II and ACCESS for ELLs are listed in Table 7. Only the correlation between Writing subtests at the Grades 3 through 5 cluster was nonsignificant, with the remaining subtest and composite scores being statistically significant. These correlations ranged from .29 to .69.

Discussion

As hypothesized, the majority of correlations among the IPT, LAS, LPTS, MAC-II, and ACCESS for ELLs were moderate to strong, suggesting that although the instruments are measuring similar constructs (i.e., ELP), the ACCESS for ELLs appears to introduce unique components that are not measured in the competing pre-NCLB ELP measures. Considering that the IPT, LAS, LPTS, and MAC-II were designed prior to NCLB accountability requirements, we would expect that there would not be a perfect correlation between scores on pre- and post-NCLB measures, and the results support this.

Although pre-NCLB measures tended to focus on social language proficiency, the passage of NCLB required that ELP measures examine more than social language proficiency and instead examine academic language proficiency, resulting in a new generation of ELP measures, such as the ACCESS for ELLs. By incorporating academic language proficiency that is linked to specific ELP standards, these measures are designed to increase the utility of obtained data, which can then be used to enhance the classroom instruction and interventions to which ELL students are exposed.

Table 6
Correlations Between Subtest Scores on the LPTS and Subtest Scores on the ACCESS for ELLs by Grade-Level Cluster and Combined Across Clusters

	ACCESS for ELLs Grade-Level Cluster											
LPTS Form	K		1–2		3–5		6–8		9–12		Combined	
ACCESS reading												
Reading												
R1, R5	.567	(146)	.768	(459)	—		—		—		.822	(605)
R2, R6	—		—		.761	(698)	—		—		.761	(698)
R3, R7	—		—		—		.658	(237)	—		.658	(237)
R4, R8	—		—		—		—		.724	(376)	.724	(376)
ACCESS writing												
Writing												
W1, W5	.479	(140)	.699	(455)	—		—		—		.759	(595)
W2, W6	—		—		.712	(691)	—		—		.712	(691)
W3, W7	—		—		—		.529	(241)	—		.529	(241)
W4, W8	—		—		—		—		.698	(268)	.698	(268)
ACCESS listening												
L/S[a]												
LS1, LS5	.455	(140)	.570	(303)	—		—		—		.640	(443)
LS2, LS6	—		—		.573	(521)	—		—		.573	(521)
LS3, LS7	—		—		—		.532	(107)	—		.532	(107)
LS4, LS8	—		—		—		—		.666	(213)	.666	(213)
ACCESS speaking												
L/S[a]												
LS1, LS5	.609	(140)	.695	(282)	—		—		—		.695	(422)
LS2, LS6	—		—		.645	(481)	—		—		.645	(481)
LS3, LS7	—		—		—		.600	(86)	—		.600	(86)
LS4, LS8	—		—		—		—		.664	(137)	.664	(137)

Note: LPTS = *Language Proficiency Test Series*; ACCESS for ELLs = *Assessing Comprehension and Communication in English State-to-State for English Language Learners*; L/S = Listening/Speaking. Values in parentheses are *n* values for the correlation analysis. All correlations were significant at $p < .001$, two-tailed.
a. The LPTS does not have separate Listening and Speaking scores.

The WIDA Consortium's ELP Standards and the ACCESS for ELLs provide examples of the new generation of standards and assessments emerging as a result of the NCLB mandate. This new generation of measures shares key characteristics that make these ELP standards and assessments more relevant for placement decisions, classroom instruction, and feedback that is both ongoing (standards used as progress benchmarks) and summative (the NCLB-mandated ELP test). The alignment of the assessment to academic goals assists teachers in teaching language through academic content, helping ELLs stay on track academically while catching up linguistically. The model language objectives, presented across a clearly defined continuum from simple to complex, provide examples for classroom teachers of how language needs can be explicitly addressed while teaching content units. Strong ELP standards should offer teachers a number of features, illustrated with the WIDA example in Table 8. The standards should be clearly related to the language of academic content classes, specifically English language arts, mathematics, and science (required under NCLB) and social studies (not federally required). They should also provide models of academic language arrayed in levels of difficulty from simple (beginners) to more complex, eventually approaching or reaching the complexity native speakers would have for specific grade levels or clusters. In the WIDA example, these are called MPIs, and each MPI includes a content stem or topic to anchor the language to the classroom, a language objective appropriate to the linguistic level, and a support idea to assist the teacher in reaching the student. The MPIs are formatted as strands from simplest (left) to most complex (right) so that teachers can easily see the progression or pathway to proficient use of the language within the respective academic context (WIDA Consortium, 2007). All of these factors are key components of intervention validity and should lead to improved student outcomes.

Table 7
Correlations Between Subtest Scores on the MAC-II and Subtest and Composite Scores on the ACCESS for ELLs by Grade-Level Cluster and Combined Across Clusters

	ACCESS for ELLs Grade-Level Cluster											
MAC-II Form	K		1–2		3–5		6–8		9–12		Combined	
ACCESS reading												
Reading												
B1	.579**	(46)	.511**	(88)	—		—		—		.607**	(134)
B2	—		.683**	(73)	.620**	(63)	—		—		.675**	(136)
B3	—		—		.633**	(125)	—		—		.633**	(125)
B4	—		—		—		.690**	(334)	—		.690**	(334)
B5	—		—		—		—		.362**	(338)	.362**	(338)
ACCESS writing												
Writing												
B1	.476*	(36)	.599**	(88)	—		—		—		.685**	(124)
B2	—		.409**	(71)	.077	(54)	—		—		.175*	(125)
B3	—		—		.577**	(125)	—		—		.577**	(125)
B4	—		—		—		.656**	(339)	—		.656**	(339)
B5	—		—		—		—		.454**	(323)	.454**	(323)
ACCESS listening												
Listening												
B1	.630**	(47)	.253*	(93)	—		—		—		.384**	(140)
B2	—		.514**	(73)	.424**	(62)	—		—		.487**	(135)
B3	—		—		.300**	(127)	—		—		.300**	(127)
B4	—		—		—		.599**	(340)	—		.599**	(340)
B5	—		—		—		—		.397**	(298)	.397**	(298)
ACCESS speaking												
Speaking												
B1	.323*	(46)	.511**	(94)	—		—		—		.403**	(140)
B2	—		.484**	(70)	.292*	(63)	—		—		.361**	(133)
B3	—		—		.330**	(126)	—		—		.330**	(126)
B4	—		—		—		.621**	(313)	—		.621**	(313)
B5	—		—		—		—		.552**	(286)	.552**	(286)
ACCESS composite												
Composite												
B1	.772**	(47)	.717**	(87)	—		—		—		.785**	(134)
B2	—		.745**	(72)	.338*	(59)	—		—		.570**	(131)
B3	—		—		.718**	(124)	—		—		.718**	(124)
B4	—		—		—		.841**	(326)	—		.841**	(326)
B5	—		—		—		—		.691**	(267)	.691**	(267)

Note: MAC-II = *Maculaitis Assessment of Competencies Test of English Language Proficiency*; ACCESS for ELLs = *Assessing Comprehension and Communication in English State-to-State for English Language Learners*. Values in parentheses are *n* values for the correlation analysis.
*$p < .05$, two-tailed. **$p < .001$, two-tailed.

Clearly, minority students—racial, ethnic, and linguistic—and children from economically disadvantaged families (which includes the vast majority of ELLs) are the students most often at risk of educational failure. ELLs are among the most disadvantaged groups in U.S. public schools today as they lack the linguistic and cultural fluency that is necessary to fully participate in the mainstream, English-speaking school setting. Their families often have limited financial resources and do not possess the social, political, and cultural capital to successfully advocate for their children. ELLs clearly face many challenges to achieving adequate academic progress.

Although our understanding of service delivery to non-ELL students has been advancing in recent years, there is a paucity of research on long-term trends in English language acquisition and its relationship with the academic achievement and social development of ELLs. With increased attention on ELP—and, in particular, academic language proficiency—the hope is that educators will be more enabled to make meaningful

Table 8
Example of the English Language Proficiency Standard 2 of the World-Class Instructional Design and Assessment Consortium: English Language Learners Communicate Information, Ideas, and Concepts Necessary for Academic Success in the Content Area of Language Arts (Framework for Formative Assessment and Instruction; Grade-Level Cluster 3–5)

	Example Topics	Level 1: Entering	Level 2: Beginning	Level 3: Developing	Level 4: Expanding	Level 5: Bridging	
Reading	Fact and opinion	Match labels or identify facts from illustrations and phrases (e.g., "I see . . ."; "There is. . .")	Identify language associated with fact in fiction or nonfiction illustrated paragraphs (e.g., "I know that . . ."; "It is true that. . .")	Identify language associated with opinion in fiction or nonfiction illustrated text (e.g., "I think that . . ."; "We believe that . . ."; "It could be . . .")	Differentiate between statements of fact and opinion found in various illustrated reading selections	Identify authors' intent associated with fact or opinion in fiction or nonfiction from grade-level text	Level 6- Reaching
	Fluency strategies	Use cues for sounding out unfamiliar words with accompanying visuals	Use visually supported context cues to derive meaning and facilitate fluency	Use punctuation cues to facilitate expression and fluency with visually supported text	Use self-monitoring and self-correcting strategies to increase fluency with visually supported text	Use strategies to adjust pace and expression while reading orally	
Writing	Editing and revising	Produce personal word or phrase lists from labeled pictures (e.g., chores or shopping) and check with a partner	Create phrases or short sentences as personal reminders from models (e.g., homework assignments) and check with a partner	Edit guided writing (e.g., for conventions and structures) based on teacher feedback	Edit and revise writing (using word processing or rubrics) based on class or peer reviews	Self-assess and revise writing (using word processing or rubrics) to produce final drafts	

differences in the education experiences of ELLs. Enhancing the instructional utility and intervention validity of ELP measures should provide educators with the necessary information to determine appropriate and powerful strategies for use with ELL students.

References

Abedi, J. (2006). Psychometric issues in the ELL assessment and special education eligibility. *Teachers College Record, 108*, 2282–2303.

Abedi, J. (2007). *English language proficiency assessment in the nation: Current status and future practice*. Davis: University of California.

August, D., & Hakuta, K. (Eds.). (1998). *Educating language minority children*. Washington, DC: National Academy Press.

Baker, S. K., Kameenui, E. J., & Simmons, D. C. (2002). Characteristics of students with diverse learning and curricular needs. In E. J. Kameenui, D. W. Carnine, R. C. Dixon, D. C. Simmons, & M. D. Coyne (Eds.), *Effective teaching strategies that accommodate diverse learners* (2nd ed., pp. 23–52). Upper Saddle River, NJ: Merrill/Prentice Hall.

Baker, S. K., Plasecia-Peinado, J., & Lezcano-Lytle, V. (1998). The use of curriculum-based measurement with language-minority

students. In M. R. Shinn (Ed.), *Advanced applications of curriculum-based measurement* (pp. 175–213). New York: Guilford.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. New York: Academic Press.

Cummins, J. (1980). The cross-lingual dimensions of language proficiency: Implications for bilingual education and the optimal age issue. *TESOL Quarterly, 14*, 175–185.

Cummins, J. (1981). The role of primary language development in promoting educational success for language minority students. In California State Department of Education (Ed.), *Schooling and language minority students: A theoretical framework* (pp. 3–49). Los Angeles: California State University, Evaluation, Dissemination and Assessment Center.

Cummins, J. C. (1984). *Bilingual and special education: Issues in assessment and pedagogy*. Austin, TX: PRO-ED.

Echevarria, J., Vogt, M. E., & Short, D. (2004). *Making content comprehensible for English language learners* (2nd ed.). Needham Heights, MA: Allyn & Bacon.

Fuchs, L. S., Fuchs, D., & Speece, D. L. (2002). Treatment validity as a unifying construct for identifying learning disabilities. *Learning Disability Quarterly, 25*, 33–45.

Gersten, R. (1999). Lost opportunities: Challenges confronting four teachers of English-language learners. *Elementary School Journal, 100*, 37–56.

Gottlieb, M. (2006). *Assessing English language learners: Bridges from language proficiency to academic achievement*. Thousand Oaks, CA: Corwin Press.

Gottlieb, M., & Kenyon, D. M. (2006). *The bridge study between tests of English language proficiency and ACCESS for ELLs* (WIDA Consortium Technical Report No. 2). Madison: Wisconsin Center for Education Research.

Gravois, T. A., & Gickling, E. E. (2002). Best practices in curriculum-based assessment. In A. Thomas & J. Grimes (Eds.), *Best practices in school psychology* (4th ed., pp. 885–898). Bethesda, MD: National Association of School Psychologists.

Gresham, F. M. (2007). Evolution of the response-to-intervention concept: Empirical foundations and recent developments. In S. R. Jimerson, M. K. Burns, & A. M. VanDerHeyden (Eds.), *Handbook of response to intervention: The science and practice of assessment and intervention* (pp. 10–24). New York: Springer.

Hayes, S. C., Nelson, R. O., & Jarrett, R. B. (1987). The treatment utility of assessment: A functional approach to evaluating assessment quality. *American Psychologist, 42*, 963–974.

Laing, S. P., & Kamhi, A. (2003). Alternative assessment of language and literacy in culturally and linguistically diverse populations. *Language, Speech, and Hearing Service in Schools, 34*, 44–55.

Lara, J. (2005). Report on an informal survey of ELL educators at the state and local levels. In C. A. Dwyer (Ed.), *Measurement and research in the accountability era* (pp. 215–224). Mahwah, NJ: Lawrence Erlbaum.

Lau v. Nichols, 414 U.S. 563 (1974).

McCardle, P., Mele-McCarthy, J., Cutting, L., Leos, K., & D-Emilio, T. (2005). Learning disabilities in English language learners: Identifying the issues. *Learning Disabilities Research and Practice, 20*, 1–5.

National Center for Education Statistics. (2007). *The condition of education 2007* (NCES 2007064). Washington, DC: U.S. Department of Education.

No Child Left Behind Act of 2001, 20 U.S.C. 70 § 6301 *et seq.* (2002).

Office for Civil Rights. (2000). *Programs for English language learners: Resource materials for planning and self-assessments.* Washington, DC: U.S. Department of Education.

Scarcella, R. (2003). *Academic English: A conceptual framework* (Technical Report 2003-1). Irvine: University of California, Linguistic Minority Research Institute.

U.S. Department of Education. (2004). *Secretary Paige announces new policies to help English language learners*. Washington, DC: Author.

U.S. Department of Education & National Institute of Child Health and Human Development. (2003). *National symposium on learning disabilities in English language learners: Symposium summary.* Washington, DC: Authors.

World-Class Instructional Design and Assessment Consortium. (2004). *WIDA English language proficiency standards for English language learners in kindergarten through Grade 12*. Madison: Wisconsin Department of Public Instruction.

World-Class Instructional Design and Assessment Consortium. (2007). *English language proficiency standards* and *Resource Guide, Pre-Kindergarten–Grade 12.* Madison: Wisconsin Center for Education Research.

Yoon, B., & Resnick, L. B. (1998). *Instructional validity, opportunity to learn, and equity: New standards examinations for the California mathematics renaissance.* Los Angeles: University of California, Los Angeles, Center for Research on Evaluation, Standards, and Student Testing.

Craig A. Albers, PhD, is an assistant professor of educational psychology in the school psychology program at the University of Wisconsin–Madison. His current interests include prevention and intervention services for students classified as English language learners and general universal screening procedures within early intervention models.

Dorry M. Kenyon, PhD, is the director of the Language Testing Division at the Center for Applied Linguistics in Washington, D.C.

Timothy J. Boals, PhD, is the executive director of the World-Class Instructional Design and Assessment (WIDA) Consortium at the Wisconsin Center for Education Research.